APPENDICES

APPENDIX A

PDBid of CD4 crystal structure

Table 17 PDBid of CD4 crystal structures reported in protein data bank (updated 6th June, 2013).

PDBid	resolu-tion(Å)	O	S	PDBid	resolu-tion(Å)	O	S	PDBid	resolu-tion(Å)	O	S
1CID	2.8	rat	monomer	1GC1	2.5	human	complex	2NY4	2	human	complex
1WBR	NMR	human	monomer	1JL4	4.3	human	complex	2NY5	2.5	human	complex
2KLU	NMR	human	monomer	1Q68	NMR	human	complex	2NY6	2.8	human	complex
1CDJ	2.5	human	monomer	1RZJ	2.2	human	complex	2QAD	3.3	human	complex
1CDU	2.7	human	monomer	1RZK	2.9	human	complex	3B71	2.98	human	complex
1CDY	2	human	monomer	2B4C	3.3	human	complex	3JWD	2.61	human	complex
1WIO	3.9	human	monomer	2JKR	2.61	human	complex	3JWO	3.51	human	complex
1WIP	4	human	monomer	2JKT	3.4	human	complex	3LQA	3.4	human	complex
1WIQ	5	human	monomer	2NXY	2	human	complex	3O2D	2.19	human	complex
3CD4	2.2	human	monomer	2NXZ	2.04	human	complex	3S4S	2.4	human	complex
1CDH	2.3	human	complex	2NY0	2.2	human	complex	3S5L	2.1	human	complex
1CDI	2.9	human	complex	2NY1	1.99	human	complex	3T0E	4	human	complex
1G9M	2.2	human	complex	2NY2	2	human	complex				
1G9N	2.9	human	complex	2NY3	2	human	complex				

O is organism, S is stoichiometry

APPENDIX B

Frequency Value in Four Criteria

Table 18 The frequency value in criterion 1.

	name	K	K	V	V	L	G	K	K	G	D	T	V	E	L	T	C	T	A	S	Q	K	K	S	I	Q
	position	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
pose	16																						1	1	4	7
	21																						1	1	3	6
	26																								1	5
	241																								2	5
	642																							1	3	6
	1302																									5
	1513																								2	6
	1128	3																								1
	1266	2																								2
	1454	3																								2
	85									1	1	2														

Table 18 (Continued) The frequency value in criterion 1.

	name	F	H	W	K	N	S	N	Q	I	K	I	L	G	N	Q	G	S	F	L	T	K	G	P	S	K
	position	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
pose	16	2	6		3			1	2	4	8				3	10	3	1		1	2	2	2	6	1	
	21	1	4		4			2	3	4	8				2	9	2	1			2	1	2	6	1	1
	26	1	3		4			2	4	4	8				2	8	2	1			1	1	2	6	1	1
	241	1	4		4			2	4	4	8				2	9	2	1	1	1	2	1	2	6	1	1
	642	2	4		5			2	5	3	8				2	8	2	1			1		2	4	1	
	1302	1	3		5			2	3	4	8				1	6	1				1		2	6	1	1
	1513	1	3		4			2	4	4	7				2	7	1	1			1	2	2	7	1	1
	1128	1	2		6	2	2	5	7	1	3															
	1266	1	2		6	1	2	5	8	1	3															
	1454	1	3		7	3	2	5	7	2	3															1
	85																		2	3	4	5	1	2	1	2

Table 18 (Continued) The frequency value in criterion 1.

	name	L	N	D	R	A	D	S	R	R	S	L	W	D	Q	G	N	F	P	L	I	I	K	N	L	K
	position	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
pose	16		1																							
	21																									
	26																									
	241																									
	642																									
	1302																									
	1513		1																							
	1128																									
	1266																									
	1454																									
	85	1	6	5	4	1	3	1	1	6	2										1		4	4		2

	name	I	E	D	S	D	T	Y	I	C	E	V	E	D	Q	K	E	E	V	Q	L	L	V
	position	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97
pose	16										2	1	4	3	1	3							
	21										3		3	2		3							
	26										3		2			3							
	241										3		2			2							
	642										3		4	3		3							
	1302										3		2			3							
	1513										3		2	2		3							
	1128						2		3		4	1	3	5	1	6	1	4		1			
	1266						2		2		4	1	3	3	1	6		2					
	1454					1	2		3		4	1	3	3	1	6		4		1			
	85																						

Table 19 The frequency value in criterion 2.

	name	K	K	V	V	L	G	K	K	G	D	T	V	E	L	T	C	T	A	S	Q	K	K	S	I	Q
	position	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
pose	16																						6	6	26	101
	21																						1	1	15	78
	26																								6	50
	241																								11	70
	642																							2	19	79
	1302																									37
	1513																								5	53
	1128	33																								13
	1266	2																								15
	1454	27																								16
	85									12	16	23														

	name	F	H	W	K	N	S	N	Q	I	K	I	L	G	N	Q	G	S	F	L	T	K	G	P	S	K
	position	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
pose	16	9	69		22			3	37	17	106				30	179	74	21		2	17	5	3	22	2	
	21	3	71		38			8	47	34	112				30	154	73	10			10	3	6	30	9	1
	26	1	63		43			16	56	63	116				15	119	41	2			8	3	5	34	15	6
	241	2	70		31			9	39	54	115				32	152	73	10	1	3	18	3	12	42	15	4
	642	3	77		51			8	44	34	111				28	152	60	6			5		6	28	4	
	1302	3	65		46			21	67	68	113				68	68	27				3		4	32	18	8
	1513	2	66		51			19	86	62	129				24	103	43	3			11	10	4	28	8	8
	1128	14	47		60	3	18	58	149	4	16															
	1266	11	43		61	1	15	46	145	6	24															
	1454	13	55		69	7	18	75	157	12	32															2
	85																		16	33	19	52	1	6	12	23

Table 19 (Continued) The frequency value in criterion 2.

	name	L	N	D	R	A	D	S	R	R	S	L	W	D	Q	G	N	F	P	L	I	I	K	N	L	K
	position	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
pose	16		2																							
	21																									
	26																									
	241																									
	642																									
	1302																									
	1513		11																							
	1128																									
	1266																									
	1454																									
	85	8	30	67	43	1	35	6	9	100	25										1		58	107		16

	name	I	E	D	S	D	T	Y	I	C	E	V	E	D	Q	K	E	E	V	Q	L	L	V
	position	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97
pose	16										31	3	59	13	1	50							
	21										42		46	11		59							
	26										42		8			59							
	241										40		15			40							
	642										53		64	37		70							
	1302										44		7			57							
	1513										49		15	4		68							
	1128						17		20		58	2	55	42	2	113	1	48		2			
	1266						6		8		54	3	50	39	3	92		23					
	1454					1	17		19		58	2	58	39	2	108		45		1			
	85																						

Table 20 The frequency value in criterion 3.

	name	K	K	V	V	L	G	K	K	G	D	T	V	E	L	T	C	T	A	S	Q	K	K	S	I	Q
	position	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
pose	16																						3	2	6	12
	21																						1	1	5	10
	26																								2	9
	241																								5	9
	642																							1	5	10
	1302																									8
	1513																								3	9
	1128	7																								7
	1266	1																								5
	1454	7																								7
	85									3	6	6														

	name	F	H	W	K	N	S	N	Q	I	K	I	L	G	N	Q	G	S	F	L	T	K	G	P	S	K
	position	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
pose	16	5	8		5			1	9	7	13				7	11	5	5		1	4	2	2	7	2	
	21	3	8		5			2	10	8	13				8	11	5	5			3	2	3	7	5	1
	26	1	8		5			4	11	8	13				7	8	5	2			2	2	3	7	5	2
	241	2	8		5			2	8	8	13				8	11	5	5	1	1	5	2	5	7	5	2
	642	3	8		5			2	8	7	12				8	9	5	4			1		3	7	2	
	1302	3	8		5			5	11	8	13				8	9	5				1		3	7	5	2
	1513	2	8		6			4	12	8	13				7	10	5	2			3	4	2	6	4	2
	1128	6	8		6	2	5	10	11	2	5															
	1266	5	8		6	1	4	10	12	2	5															
	1454	6	8		6	3	5	10	12	4	8															1
	85																		6	6	6	13	1	3	4	6

Table 20 (Continued) The frequency value in criterion 3.

	name	L	N	D	R	A	D	S	R	R	S	L	W	D	Q	G	N	F	P	L	I	I	K	N	L	K
	position	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
pose	16		1																							
	21																									
	26																									
	241																									
	642																									
	1302																									
	1513		5																							
	1128																									
	1266																									
	1454																									
	85	4	10	9	12	1	4	2	4	12	5										1		9	9		5

	name	I	E	D	S	D	T	Y	I	C	E	V	E	D	Q	K	E	E	V	Q	L	L	V
	position	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97
pose	16										8	2	10	6	1	6							
	21										8		10	4		6							
	26										8		4			6							
	241										8		7			5							
	642										8		10	7		7							
	1302										8		3			6							
	1513										8		8	3		6							
	1128						4		4		8	2	10	8	2	11	1	5		1			
	1266						4		3		8	3	9	8	2	8		4					
	1454					1	5		4		8	2	10	8	2	11		4		1			
	85																						

Table 21 The frequency value in criterion 4.

	name	K	K	V	V	L	G	K	K	G	D	T	V	E	L	T	C	T	A	S	Q	K	K	S	I	Q
	position	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
pose	16																						23.1	25	66.7	100
	21																						7.69	12.5	55.6	83.3
	26																								22.2	75
	241																								55.6	75
	642																							12.5	55.6	83.3
	1302																									66.7
	1513																								33.3	75
	1128	53.8																								58.3
	1266	7.69																								41.7
	1454	53.8																								58.3
	85									60	66.7	66.7														

	name	F	H	W	K	N	S	N	Q	I	K	I	L	G	N	Q	G	S	F	L	T	K	G	P	S	K
	position	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
pose	16	41.7	66.7		38.5			9.09	75	77.8	100				63.6	91.7	100	62.5		11.1	44.4	15.4	40	100	25	
	21	25	66.7		38.5			18.2	83.3	88.9	100				72.7	91.7	100	62.5			33.3	15.4	60	100	62.5	7.69
	26	8.33	66.7		38.5			36.4	91.7	88.9	100				63.6	66.7	100	25			22.2	15.4	60	100	62.5	15.4
	241	16.7	66.7		38.5			18.2	66.7	88.9	100				72.7	91.7	100	62.5	8.33	11.1	55.6	15.4	100	100	62.5	15.4
	642	25	66.7		38.5			18.2	66.7	77.8	92.3				72.7	75	100	50			11.1		60	100	25	
	1302	25	66.7		38.5			45.5	91.7	89	100				72.7	75	100				11.1		60	100	62.5	15.4
	1513	16.7	66.7		46.2			36.4	100	89	100				63.6	83.3	100	25			33.3	30.8	40	85.7	50	15.4
	1128	50	66.7		46.2	18.2	62.5	90.9	91.7	22	38.5															
	1266	41.7	66.7		46.2	9.09	50	90.9	100	22	38.5															
	1454	50	66.7		46.2	27.3	62.5	90.9	100	44	61.5															7.69
	85																		50	66.7	66.7	100	20	42.9	50	46.2

Table 21 (Continued) The frequency value in criterion 4.

	name	L	N	D	R	A	D	S	R	R	S	L	W	D	Q	G	N	F	P	L	I	I	K	N	L	K
	position	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
pose	16		9																							
	21																									
	26																									
	241																									
	642																									
	1302																									
	1513		45.5																							
	1128																									
	1266																									
	1454																									
	85	44.4	90.9	100	35294	16.7	44.4	25	23.5	35294	62.5										11.1		69.2	81.8		38.5

	name	I	E	D	S	D	T	Y	I	C	E	V	E	D	Q	K	E	E	V	Q	L	L	V
	position	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97
pose	16										80	25	100	66.7	8.33	46.2							
	21										80		100	44.4		46.2							
	26										80		40			46.2							
	241										80		70			38.5							
	642										80		100	77.8		53.8							
	1302										80		30			46.2							
	1513										80		80	33.3		46.2							
	1128						44.4		44.4		80	25	100	88.9	16.7	84.6	10	50		8.33			
	1266						44.4		33.3		80	37.5	90	88.9	16.7	61.5		40					
	1454					11.1	55.6		44.4		80	25	100	88.9	16.7	84.6		40		8.33			
	85																						

APPENDIX C

Frequency Value of Considered CD4 Residues in 11 Poses

Table 22 The raw frequency value of considered CD4 residues in 11 poses.

A. Pose number 16 (p16)

considered CD4 residues of p16														
position		**24**	**25**	**27**	**33**	**34**	**35**	**39**	**40**	**41**	**48**	**85**	**87**	**88**	**90**
criteria	**1**	4	7	6	2	4	8	3	10	3	6	2	4	3	3
	2	26	101	69	37	17	106	30	179	74	22	31	59	13	50
	3	6	12	8	9	7	13	7	11	5	7	8	10	6	6
	4	66.67	100	66.67	75	77.78	100	63.64	91.67	100	100	80	100	66.67	46.15
	5	0	1	0	0	0	0	0	2	0	0	0	0	0	1

Table 22 (Continued) The raw frequency value of considered CD4 residues in 11 poses.

B. Pose number 21 (p21)

considered CD4 residues of p21													
position		**25**	**27**	**29**	**33**	**34**	**35**	**39**	**40**	**41**	**48**	**85**	**87**	**90**
criteria	**1**	6	4	4	3	4	8	2	9	2	6	3	3	3
	2	78	71	38	47	34	112	30	154	73	30	42	46	59
	3	10	8	5	10	8	13	8	11	5	7	8	10	6
	4	83.33	66.67	38.46	83.33	88.89	100	72.73	91.67	100	100	80	100	46.15
	5	1	0	0	0	0	1	0	2	0	0	0	0	1

C. Pose number 24 (p26)

considered CD4 residues of p26													
position		**25**	**27**	**29**	**33**	**34**	**35**	**39**	**40**	**41**	**48**	**85**	**90**
criteria	**1**	5	3	4	4	4	8	2	8	2	6	3	3
	2	50	63	43	56	63	116	15	119	41	34	42	59
	3	9	8	5	11	8	13	7	8	5	7	8	6
	4	75	66.67	38.46	91.67	88.89	100	63.64	66.67	100	100	80	46.15
	5	1	0	0	0	0	1	0	2	0	0	0	0

Table 22 (Continued) The raw frequency value of considered CD4 residues in 11 poses.

D. Pose number 241 (p241)

considered CD4 residues of p241														
position		**25**	**27**	**29**	**33**	**34**	**35**	**39**	**40**	**41**	**47**	**48**	**85**	**87**	**90**
criteria	**1**	5	4	4	4	4	8	2	9	2	2	6	3	2	2
	2	70	70	31	39	54	115	32	152	73	12	42	40	15	40
	3	9	8	5	8	8	13	8	11	5	5	7	8	7	5
	4	75	66.67	38.46	66.67	88.89	100	72.73	91.67	100	100	100	80	70	38.46
	5	1	0	0	0	0	0	0	1	0	0	0	0	0	0

E. Pose number 642 (p642)

considered CD4 residues of p642															
position		**25**	**27**	**29**	**33**	**34**	**35**	**39**	**40**	**41**	**48**	**85**	**87**	**88**	**90**
criteria	**1**	6	4	5	5	3	8	2	8	2	4	3	4	3	3
	2	79	77	51	44	34	111	28	152	60	28	53	64	37	70
	3	10	8	5	8	7	12	8	9	5	7	8	10	7	7
	4	83.33	66.67	38.46	66.67	77.78	92.31	72.73	75	100	100	80	100	77.78	53.85
	5	1	0	0	0	0	1	0	1	0	0	0	0	0	1

Table 22 (Continued) The raw frequency value of considered CD4 residues in 11 poses.

F. Pose number 1302 (p1302)

considered CD4 residues of p1302													
		25	**27**	**29**	**33**	**34**	**35**	**39**	**40**	**41**	**48**	**85**	**90**
criteria	**1**	5	3	5	3	4	8	1	6	1	6	3	3
	2	37	65	46	67	68	113	68	68	27	32	44	57
	3	8	8	5	11	8	13	8	9	5	7	8	6
	4	66.67	66.67	38.46	91.67	89	100	72.73	75	100	100	80	46.15
	5	0	0	0	0	0	1	0	0	0	0	0	0

G. Pose number 1513 (p1513)

considered CD4 residues of p1513														
position		**25**	**27**	**29**	**33**	**34**	**35**	**39**	**40**	**41**	**48**	**85**	**87**	**90**
criteria	**1**	6	3	4	4	4	7	2	7	1	7	3	2	3
	2	53	66	51	86	62	129	24	103	43	28	49	15	68
	3	9	8	6	12	8	13	7	10	5	6	8	8	6
	4	75	66.67	46.15	100	89	100	63.64	83.33	100	85.71	80	80	46.15
	5	1	0	2	0	0	2	0	0	0	0	0	0	1

Table 22 (Continued) The raw frequency value of considered CD4 residues in 11 poses.

H. Pose number 1128 (p1128)

considered CD4 residues of p1128													
position		**1**	**25**	**27**	**29**	**31**	**32**	**33**	**85**	**87**	**88**	**90**	**92**
criteria	**1**	3	1	2	6	2	5	7	4	3	5	6	4
	2	33	13	47	60	18	58	149	58	55	42	113	48
	3	7	7	8	6	5	10	11	8	10	8	11	5
	4	53.85	58.33	66.67	46.15	62.5	90.91	91.67	80	100	88.89	84.62	50
	5	0	0	0	1	0	1	1	0	0	0	1	1

I. Pose number 1266 (p1266)

considered CD4 residues of p1266													
position		**27**	**29**	**31**	**32**	**33**	**35**	**81**	**85**	**87**	**88**	**90**	**92**
criteria	**1**	2	6	2	5	8	3	2	4	3	3	6	2
	2	43	61	15	46	145	24	6	54	50	39	92	23
	3	8	6	4	10	12	5	4	8	9	8	8	4
	4	66.67	46.15	50	90.91	100	38.46	44.44	80	90	88.89	61.54	40
	5	0	1	0	0	0	0	0	0	0	0	2	0

Table 22 (Continued) The raw frequency value of considered CD4 residues in 11 poses.

J. Pose number 1454

considered CD4 residues of p1454													
position		**25**	**27**	**29**	**31**	**32**	**33**	**35**	**85**	**87**	**88**	**90**	**92**
criteria	**1**	2	3	7	2	5	7	3	4	3	3	6	4
	2	16	55	69	18	75	157	32	58	58	39	108	45
	3	7	8	6	5	10	12	8	8	10	8	11	4
	4	58.33	66.67	46.15	62.5	90.91	100	61.54	80	100	88.89	84.62	40
	5	0	0	1	0	0	2	0	0	0	0	2	1

K. Pose number 85 (p85)

considered CD4 residues of p85												
position		**44**	**45**	**46**	**52**	**53**	**54**	**56**	**59**	**60**	**72**	**73**
criteria	**1**	3	4	5	6	5	4	3	6	2	4	4
	2	33	19	52	30	67	43	35	100	25	58	107
	3	6	6	13	10	9	12	4	12	5	9	9
	4	66.67	66.67	100	90.91	100	70.59	44.44	70.59	62.5	69.23	81.82
	5	0	0	0	0	0	0	0	1	0	0	0

Table 23 The normalized value of considered CD4 residues in 5 criteria of 11 poses.

A. Pose number 16 (p16)

considered CD4 residues of p16														
position		**24**	**25**	**27**	**33**	**34**	**35**	**39**	**40**	**41**	**48**	**85**	**87**	**88**	**90**
criteria	**1**	-0.27	0.98	0.56	-1.1	-0.27	1.4	-0.68	2.23	-0.68	0.56	-1.1	-0.27	-0.68	-0.68
	2	-0.7	0.94	0.24	-0.46	-0.9	1.05	-0.62	2.64	0.35	-0.79	-0.59	0.02	-0.99	-0.18
	3	-0.9	1.54	-0.09	0.32	-0.49	1.95	-0.49	1.13	-1.31	-0.49	-0.09	0.73	-0.9	-0.9
	4	-0.81	1.07	-0.81	-0.34	-0.18	1.07	-0.98	0.6	1.07	1.07	-0.06	1.07	-0.81	-1.96
	5	-0.47	1.17	-0.47	-0.47	-0.47	-0.47	-0.47	2.8	-0.47	-0.47	-0.47	-0.47	-0.47	1.17

B. Pose number 21 (p21)

considered CD4 residues of p21														
position		**25**	**27**	**29**	**33**	**34**	**35**	**39**	**40**	**41**	**48**	**85**	**87**	**90**
criteria	**1**	0.73	-0.17	-0.17	-0.62	-0.17	1.63	-1.07	2.08	-1.07	0.73	-0.62	-0.62	-0.62
	2	0.43	0.23	-0.68	-0.43	-0.79	1.36	-0.9	2.53	0.29	-0.9	-0.57	-0.46	-0.1
	3	0.68	-0.16	-1.43	0.68	-0.16	1.95	-0.16	1.11	-1.43	-0.59	-0.16	0.68	-1.01
	4	0.12	-0.7	-2.09	0.12	0.4	0.94	-0.4	0.53	0.94	0.94	-0.04	0.94	-1.71
	5	0.95	-0.59	-0.59	-0.59	-0.59	0.95	-0.59	2.48	-0.59	-0.59	-0.59	-0.59	0.95

Table 23 (Continued) The normalized value of considered CD4 residues in 5 criteria of 11 poses.

C. Pose number 26 (p26)

considered CD4 residues of p26													
position		**25**	**27**	**29**	**33**	**34**	**35**	**39**	**40**	**41**	**48**	**85**	**90**
criteria	**1**	0.32	-0.65	-0.16	-0.16	-0.16	1.78	-1.13	1.78	-1.13	0.81	-0.65	-0.65
	2	-0.27	0.15	-0.5	-0.08	0.15	1.87	-1.41	1.97	-0.57	-0.79	-0.53	0.02
	3	0.47	0.04	-1.26	1.33	0.04	2.2	-0.4	0.04	-1.26	-0.4	0.04	-0.83
	4	-0.07	-0.47	-1.82	0.73	0.6	1.13	-0.61	-0.47	1.13	1.13	0.17	-1.45
	5	1.02	-0.51	-0.51	-0.51	-0.51	1.02	-0.51	2.56	-0.51	-0.51	-0.51	-0.51

D. Pose number 241 (p241)

considered CD4 residues of p241															
position		**25**	**27**	**29**	**33**	**34**	**35**	**39**	**40**	**41**	**47**	**48**	**85**	**87**	**90**
criteria	**1**	0.41	-0.03	-0.03	-0.03	-0.03	1.73	-0.91	2.17	-0.91	-0.91	0.85	-0.47	-0.91	-0.91
	2	0.36	0.36	-0.65	-0.44	-0.05	1.53	-0.63	2.5	0.44	-1.15	-0.37	-0.42	-1.07	-0.42
	3	0.58	0.15	-1.13	0.15	0.15	2.29	0.15	1.43	-1.13	-1.13	-0.27	0.15	-0.27	-1.13
	4	-0.13	-0.53	-1.87	-0.53	0.53	1.06	-0.24	0.66	1.06	1.06	1.06	0.11	-0.37	-1.87
	5	2.36	-0.39	-0.39	-0.39	-0.39	-0.39	-0.39	2.36	-0.39	-0.39	-0.39	-0.39	-0.39	-0.39

Table 23 (Continued) The normalized value of considered CD4 residues in 5 criteria of 11 poses.

E. Pose number 642 (p642)

considered CD4 residues of p642														
position		**25**	**27**	**29**	**33**	**34**	**35**	**39**	**40**	**41**	**48**	**85**	**87**	**88**	**90**
criteria	**1**	0.88	-0.15	0.37	0.37	-0.66	1.92	-1.18	1.92	-1.18	-0.15	-0.66	-0.15	-0.66	-0.66
	2	0.45	0.4	-0.36	-0.57	-0.86	1.39	-1.03	2.58	-0.1	-1.03	-0.3	0.02	-0.77	0.19
	3	1.09	0.04	-1.54	0.04	-0.49	2.14	0.04	0.56	-1.54	-0.49	0.04	1.09	-0.49	-0.49
	4	0.33	-0.61	-2.19	-0.61	0.02	0.83	-0.27	-0.14	1.26	1.26	0.14	1.26	0.02	-1.33
	5	1.52	-0.61	-0.61	-0.61	-0.61	1.52	-0.61	1.52	-0.61	-0.61	-0.61	-0.61	-0.61	1.52

F. Pose number 1302 (p1302)

considered CD4 residues of p1302													
position		**25**	**27**	**29**	**33**	**34**	**35**	**39**	**40**	**41**	**48**	**85**	**90**
criteria	**1**	0.48	-0.48	0.48	-0.48	0	1.91	-1.44	0.96	-1.44	0.96	-0.48	-0.48
	2	-0.89	0.32	-0.5	0.4	0.45	2.39	0.45	0.45	-1.33	-1.11	-0.59	-0.03
	3	0	0	-1.31	1.31	0	2.18	0	0.44	-1.31	-0.44	0	-0.87
	4	-0.51	-0.51	-1.89	0.71	0.58	1.11	-0.22	-0.11	1.11	1.11	0.14	-1.51
	5	-0.29	-0.29	-0.29	-0.29	-0.29	3.18	-0.29	-0.29	-0.29	-0.29	-0.29	-0.29

Table 23 (Continued) The normalized value of considered CD4 residues in 5 criteria of 11 poses.

G. Pose number 1513 (p1513)

considered CD4 residues of p1513													
position		**25**	**27**	**29**	**33**	**34**	**35**	**39**	**40**	**41**	**48**	**85**	**87**	**90**
criteria	**1**	0.93	-0.52	-0.04	-0.04	-0.04	1.42	-1.01	1.42	-1.49	1.42	-0.52	-1.01	-0.52
	2	-0.21	0.19	-0.27	0.82	0.07	2.16	-1.12	1.35	-0.52	-0.99	-0.34	-1.4	0.26
	3	0.36	-0.06	-0.91	1.62	-0.06	2.04	-0.49	0.78	-1.33	-0.91	-0.06	-0.06	-0.91
	4	-0.17	-0.62	-1.74	1.19	0.59	1.19	-0.79	0.28	1.19	0.41	0.1	0.1	-1.74
	5	0.69	-0.59	1.98	-0.59	-0.59	1.98	-0.59	-0.59	-0.59	-0.59	-0.59	-0.59	0.69

H. Pose number 1128 (p1128)

considered CD4 residues of p1128													
position		**1**	**25**	**27**	**29**	**31**	**32**	**33**	**85**	**87**	**88**	**90**	**92**
criteria	**1**	-0.54	-1.61	-1.08	1.08	-1.08	0.54	1.61	0	-0.54	0.54	1.08	0
	2	-0.65	-1.17	-0.28	0.06	-1.04	0	2.39	0	-0.07	-0.41	1.44	-0.26
	3	-0.47	-0.47	0	-0.94	-1.41	0.94	1.41	0	0.94	0	1.41	-1.41
	4	-1.02	-0.78	-0.33	-1.43	-0.55	0.97	1.01	0.39	1.46	0.86	0.63	-1.22
	5	-0.81	-0.81	-0.81	1.13	-0.81	1.13	1.13	-0.81	-0.81	-0.81	1.13	1.13

Table 23 (Continued) The normalized value of considered CD4 residues in 5 criteria of 11 poses.

I. Pose number 1266 (p1266)

considered CD4 residues of p1266													
position		**27**	**29**	**31**	**32**	**33**	**35**	**81**	**85**	**87**	**88**	**90**	**92**
criteria	1	-0.92	1.09	-0.92	0.59	2.09	-0.42	-0.92	0.08	-0.42	-0.42	1.09	-0.92
	2	-0.18	0.3	-0.92	-0.1	2.52	-0.68	-1.16	0.11	0	-0.29	1.12	-0.71
	3	0.32	-0.45	-1.22	1.09	1.87	-0.84	-1.22	0.32	0.71	0.32	0.32	-1.22
	4	0.01	-0.9	-0.73	1.08	1.48	-1.23	-0.97	0.6	1.04	0.99	-0.22	-1.17
	5	-0.4	1.21	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	2.82	-0.4

J. Pose number 1454 (p1454)

considered CD4 residues of p1454													
position		**25**	**27**	**29**	**31**	**32**	**33**	**35**	**85**	**87**	**88**	**90**	**92**
criteria	**1**	-1.17	-0.61	1.64	-1.17	0.51	1.64	-0.61	-0.05	-0.61	-0.61	1.08	-0.05
	2	-1.13	-0.15	0.21	-1.08	0.36	2.43	-0.73	-0.07	-0.07	-0.55	1.19	-0.4
	3	-0.45	-0.03	-0.87	-1.29	0.8	1.64	-0.03	-0.03	0.8	-0.03	1.22	-1.71
	4	-0.74	-0.33	-1.34	-0.53	0.87	1.32	-0.58	0.33	1.32	0.77	0.56	-1.64
	5	-0.63	-0.63	0.63	-0.63	-0.63	1.88	-0.63	-0.63	-0.63	-0.63	1.88	0.63

Table 23 (Continued) The normalized value of considered CD4 residues in 5 criteria of 11 poses.

K. Pose number 85 (p85)

considered CD4 residues of p85												
position		**44**	**45**	**46**	**52**	**53**	**54**	**56**	**59**	**60**	**72**	**73**
criteria	**1**	-0.95	-0.15	0.65	1.45	0.65	-0.15	-0.95	1.45	-1.74	-0.15	-0.15
	2	-0.64	-1.11	0.01	-0.74	0.52	-0.3	-0.57	1.64	-0.91	0.21	1.88
	3	-0.87	-0.87	1.43	0.45	0.12	1.11	-1.52	1.11	-1.2	0.12	0.12
	4	-0.49	-0.49	1.49	0.95	1.49	-0.25	-1.8	-0.25	-0.73	-0.33	0.41
	5	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	3.02	-0.3	-0.3	-0.3

APPENDIX D

Prediction Result from HotPOINT and HSPred Software

Table 24 The hot spots and non-hot spots with their scoring value of HotPOINT prediction.

A. Pose number 16 (p16)

prediction results of p16					
residue number	**residue name**	**RelComp ASA**	**RelMonomer ASA**	**Potential**	**prediction**
24	I	1.75	27.04	18.52	H
25	Q	1.03	56.38	31.05	H
27	H	0	32.62	19	H
33	Q	45.18	68.83	19.69	NH
34	I	13.37	38.73	13	NH
35	K	3.87	47.3	21.21	H
39	N	2.45	11.53	21.59	H
40	Q	0.03	75.38	37.67	H
41	G	7.44	74.97	11.57	NH
42	S	53.6	64.6	7.34	NH
45	T	7.16	39.23	18.35	H
48	P	27.23	92.84	14.34	NH
85	E	7.69	18.86	20.53	H
87	E	11.99	44.33	17.17	NH
90	K	35.4	53.5	8.73	NH

Table 24 (Continued) The hot spots and non-hot spots with their scoring value of HotPOINT prediction.

B. Pose number 21 (p21)

prediction results of p21					
residue number	**residue name**	**RelComp ASA**	**RelMonomer ASA**	**Potential**	**prediction**
24	I	6.67	31.21	18.52	H
25	Q	2.12	51.75	25.69	H
27	H	0	29.36	21.47	H
33	Q	39.95	65.79	19.69	NH
34	I	12.38	40.81	15.88	NH
35	K	5.55	50.05	21.21	H
39	N	3.29	10.57	19.53	H
40	Q	1.42	77.65	37.67	H
41	G	16.39	78.33	11.57	NH
45	T	12.34	43.38	15.02	NH
47	G	3.92	19.7	7.42	NH
48	P	23.21	89.99	14.34	NH
49	S	23.3	32.12	9.43	NH
85	E	4.58	19.31	20.53	H
87	E	14.7	46.08	17.17	NH
90	K	33.89	52.9	11.2	NH

Table 24 (Continued) The hot spots and non-hot spots with their scoring value of HotPOINT prediction.

C. Pose number 26 (p26)

prediction results of p26					
residue number	**residue name**	**RelComp ASA**	**RelMonomer ASA**	**Potential**	**prediction**
24	I	19.48	32.34	17.65	NH
25	Q	3.91	48.04	25.69	H
27	H	0	26.68	21.47	H
29	K	0.53	7.8	19.51	H
32	N	70.48	79.67	4.13	NH
33	Q	36.93	63.38	19.69	NH
34	I	8.3	41.45	15.88	NH
35	K	8.51	52.24	21.48	H
40	Q	5.71	80.46	33.02	H
41	G	31.6	77.04	6.01	NH
45	T	15.73	46.09	13.38	NH
47	G	5.78	21.97	7.42	NH
48	P	20.81	90.59	14.34	NH
49	S	20.4	32.13	13.44	NH
85	E	4.27	18.72	20.53	H
87	E	22.65	45.8	14.32	NH
90	K	34.45	52.83	11.2	NH

Table 24 (Continued) The hot spots and non-hot spots with their scoring value of HotPOINT prediction.

D. Pose number 241 (p241)

prediction results of p241					
residue number	**residue name**	**RelComp ASA**	**RelMonomer ASA**	**Potential**	**prediction**
24	I	9.64	30.56	18.52	H
25	Q	2.08	51.79	25.69	H
27	H	0	29.55	19	H
33	Q	41.76	65.51	17.13	NH
34	I	12.35	42.19	18.76	H
35	K	6.51	53.85	24.44	H
39	N	2.92	13.09	17.7	NH
40	Q	0.25	77.22	37.67	H
41	G	13.17	78.85	11.57	NH
45	T	7.8	44.46	18.35	H
47	G	2.11	21.74	8.74	NH
48	P	15.42	86.61	14.34	NH
49	S	19.97	33.03	16.39	NH
85	E	5.16	19.42	20.53	H
87	E	20.15	46.01	14.32	NH
90	K	38.41	52.19	11.2	NH

Table 24 (Continued) The hot spots and non-hot spots with their scoring value of HotPOINT prediction.

E. Pose number 642 (p642)

prediction results of p642					
residue number	**residue name**	**RelComp ASA**	**RelMonomer ASA**	**Potential**	**prediction**
24	I	4.84	31.02	18.52	H
25	Q	2.21	51.34	31.05	H
27	H	0.09	29.34	21.47	H
29	K	0.51	7.99	16.68	NH
33	Q	39.18	66.23	17.13	NH
34	I	17.2	41.15	15.88	NH
35	K	8.24	51.33	19.57	H
39	N	4.31	10.07	19.53	H
40	Q	2	81.42	30	H
41	G	21.99	77.05	9.8	NH
45	T	25.38	52.1	13.38	NH
47	G	6.03	21.34	7.42	NH
48	P	34.09	89.62	14.34	NH
85	E	2.49	17.71	20.53	H
87	E	12.23	46.33	19.86	H
88	D	56.87	98.77	11.21	NH
90	K	32.67	52.4	11.2	NH

Table 24 (Continued) The hot spots and non-hot spots with their scoring value of HotPOINT prediction.

F. Pose number 1302 (p1302)

prediction results of p1302					
residue number	**residue name**	**RelComp ASA**	**RelMonomer ASA**	**Potential**	**prediction**
25	Q	7.61	50.31	25.69	H
27	H	0.11	28.19	21.47	H
29	K	0.41	7.94	19.51	H
32	N	68.57	80.73	4.13	NH
33	Q	35.83	61.34	19.69	NH
34	I	5.64	39.49	18.76	H
35	K	11.11	52.14	21.48	H
40	Q	19.29	84.19	30.59	H
41	G	39.99	75.18	3.77	NH
47	G	9.31	20.31	7.42	NH
48	P	21.34	93.12	10.88	NH
49	S	20.03	30.83	13.44	NH
85	E	4.31	18.64	20.53	H
87	E	32.45	45.32	10.27	NH
90	K	33.62	52.11	11.2	NH

Table 24 (Continued) The hot spots and non-hot spots with their scoring value of HotPOINT prediction.

G. Pose number 1513 (p1513)

prediction results of p1513					
residue number	**residue name**	**RelComp ASA**	**RelMonomer ASA**	**Potential**	**prediction**
25	Q	3.47	47.36	25.69	H
27	H	0	24.95	16.53	NH
29	K	0.03	10.66	17.29	NH
32	N	68.48	82.38	4.13	NH
33	Q	36.51	60.29	19.69	NH
34	I	1.85	35.68	18.76	H
35	K	7.53	55.02	21.48	H
39	N	5.02	12.21	19.06	H
40	Q	10.86	79.69	35.24	H
41	G	32.43	73.25	3.77	NH
45	T	17.61	49.39	13.38	NH
48	P	7.15	93.69	13.8	NH
49	S	22.57	32.87	13.44	NH
85	E	4.76	17.86	23.38	H
87	E	20.52	44.31	14.32	NH
90	K	32.29	52.61	12.52	NH

Table 24 (Continued) The hot spots and non-hot spots with their scoring value of HotPOINT prediction.

H. Pose number 1128 (p1128)

prediction results of p1128					
residue number	**residue name**	**RelComp ASA**	**RelMonomer ASA**	**Potential**	**prediction**
25	Q	33.61	51.3	18.07	NH
26	F	0.85	2.89	40.62	H
27	H	2.87	29.99	17.85	NH
29	K	0.38	7.88	21.8	H
31	S	14.82	40.7	16.22	NH
32	N	8.96	77.87	16	NH
33	Q	7.09	66.84	40.2	H
81	T	6.68	31.87	14.72	NH
83	I	0	13.32	20.3	H
85	E	0	22	27.94	H
87	E	17.16	39.66	12.61	NH
88	D	73.34	101.9	6.66	NH
90	K	20.47	50.54	13.96	NH
92	E	2.61	30.95	23.72	H

Table 24 (Continued) The hot spots and non-hot spots with their scoring value of HotPOINT prediction.

I. Pose number 1266 (p1266)

prediction results of p1266					
residue number	**residue name**	**RelComp ASA**	**RelMonomer ASA**	**Potential**	**prediction**
25	Q	29.89	50.82	18.07	NH
26	F	0.69	2.64	40.62	H
27	H	1.88	29.05	17.85	NH
29	K	0.07	7.23	21.8	H
31	S	14.21	41.25	16.22	NH
32	N	17.7	78.18	16	NH
33	Q	5.6	67.6	31.9	H
34	I	22.64	39.27	15.88	NH
35	K	30.14	51.7	14.4	NH
81	T	17.47	36.61	11.85	NH
83	I	0	11.97	15.65	NH
85	E	0	20.86	27.94	H
87	E	19.72	42.2	12.61	NH
88	D	79.18	102.29	2.41	NH
90	K	23.98	51.46	16.18	NH

Table 24 (Continued) The hot spots and non-hot spots with their scoring value of HotPOINT prediction.

J. Pose number 1454 (p1454)

prediction results of p1454					
residue number	**residue name**	**RelComp ASA**	**RelMonomer ASA**	**Potential**	**prediction**
25	Q	28.96	50.38	18.07	NH
26	F	0.8	2.65	40.62	H
27	H	0.8	29.24	17.85	NH
29	K	0.27	7.83	21.8	H
31	S	13.37	37.21	19.69	H
32	N	4.16	78.91	16	NH
33	Q	3.31	65.3	43.85	H
34	I	19.5	38.8	15.88	NH
35	K	28.31	52.94	14.4	NH
81	T	7.37	33.86	14.72	NH
83	I	0	12.4	20.3	H
85	E	0	20.67	27.94	H
87	E	17.24	40.69	12.61	NH
88	D	78.7	101.96	4.82	NH
90	K	20.58	51.09	17.2	NH
92	E	1.82	31.67	23.72	H

Table 24 (Continued) The hot spots and non-hot spots with their scoring value of HotPOINT prediction.

K. Pose number 85(p85)

prediction results of p85					
residue number	**residue name**	**RelComp ASA**	**RelMonomer ASA**	**Potential**	**prediction**
9	G	39.46	60.17	9.19	NH
10	D	27.74	30.47	4.16	NH
11	T	42.4	66.53	7.5	NH
44	L	1.81	21.21	41.02	H
45	T	19.35	54.58	10.74	NH
46	K	1.8	25.99	19.1	H
49	S	18.87	30.11	9.43	NH
50	K	32.28	46.35	16.88	NH
52	N	7.17	37.9	19.89	H
53	D	3.18	68.63	21.63	H
54	R	0.15	17.93	18.22	H
56	D	2.12	30.23	22.99	H
57	S	1.24	9.74	20.14	H
59	R	21.55	57.2	15.82	NH
60	S	40.45	80.92	6.34	NH
70	I	2.87	7.32	26.17	H
72	K	2.47	30.78	11.97	NH
73	N	9.32	51.19	15.24	NH

Table 25 The hot spots and non-hot spots with their scoring value of HSPred prediction.

p16			p21			p26		
position	**name**	**ΔΔG**	**position**	**name**	**ΔΔG**	**position**	**name**	**ΔΔG**
A23	S	-1.49	A23	S	-1.558	A24	I	-1.292
A24	I	-0.607	A24	I	-0.973	A25	Q	-0.425
A25	Q	0.666	A25	Q	0.443	A26	F	-0.465
A26	F	0.093	A26	F	0.165	A27	H	0.588
A27	H	0.634	A27	H	0.713	A29	K	0.362
A29	K	-0.14	A29	K	0.206	A32	N	-1.351
A32	N	-1.554	A32	N	-1.445	A33	Q	-0.172
A33	Q	-0.693	A33	Q	-0.693	A34	I	-0.223
A34	I	-1.024	A34	I	-0.464	A35	K	1.003
A35	K	0.492	A35	K	0.924	A39	N	-0.878
A39	N	-0.357	A39	N	-0.569	A40	Q	0.115
A40	Q	0.742	A40	Q	0.55	A45	T	-1.221
A42	S	-1.379	A42	S	-1.535	A46	K	-1.38
A44	L	-1.462	A45	T	-1.077	A49	S	-1.322
A45	T	-0.967	A46	K	-1.378	A50	K	-1.481
A46	K	-1.379	A49	S	-1.437	A85	E	-0.52
A49	S	-1.568	A85	E	-0.433	A87	E	-1.566
A85	E	-0.155	A87	E	-1.036	A90	K	-0.859
A86	V	-1.052	A88	D	-1.125			
A87	E	-0.747	A90	K	-0.783			
A88	D	-1.017						
A90	K	-0.744						

Table 25 (Continued) The hot spots and non-hot spots with their scoring value of HSPred prediction.

p241			p642			p1302		
position	**name**	**ΔΔG**	**position**	**name**	**ΔΔG**	**position**	**name**	**ΔΔG**
A24	I	-1.094	A23	S	-1.572	A25	Q	-0.848
A25	Q	0.032	A24	I	-0.899	A26	F	-0.542
A26	F	-0.256	A25	Q	0.487	A27	H	0.545
A27	H	0.319	A26	F	-0.114	A29	K	0.518
A29	K	-0.16	A27	H	0.865	A32	N	-1.22
A32	N	-1.45	A29	K	0.408	A33	Q	-0.002
A33	Q	-0.77	A32	N	-1.503	A34	I	-0.083
A34	I	-0.655	A33	Q	-0.262	A35	K	0.941
A35	K	0.453	A34	I	-0.467	A39	N	-1.347
A39	N	-0.548	A35	K	0.903	A40	Q	-0.723
A40	Q	0.565	A39	N	-0.702	A45	T	-1.564
A42	S	-1.516	A40	Q	0.289	A49	S	-1.231
A43	F	-1.12	A42	S	-1.576	A50	K	-1.412
A44	L	-1.445	A45	T	-1.305	A85	E	-0.97
A45	T	-0.989	A49	S	-1.459	A87	E	-1.569
A46	K	-1.291	A85	E	-0.154	A90	K	-0.813
A49	S	-1.267	A87	E	-0.789			
A50	K	-1.462	A88	D	-0.689			
A85	E	-0.428	A90	K	-0.584			
A87	E	-1.025						
A90	K	-1.041						

Table 25 (Continued) The hot spots and non-hot spots with their scoring value of HSPred prediction.

p1513			p1128			p1266		
position	**name**	**ΔΔG**	**position**	**name**	**ΔΔG**	**position**	**name**	**ΔΔG**
A24	I	-1.249	A1	K	-0.915	A25	Q	-1.151
A25	Q	-0.305	A25	Q	-1.122	A26	F	-0.52
A26	F	-0.354	A26	F	-0.44	A27	H	-0.101
A27	H	0.805	A27	H	0.057	A29	K	0.343
A29	K	0.862	A29	K	0.847	A30	N	-1.156
A32	N	-1.043	A30	N	-1.137	A31	S	-1.295
A33	Q	-0.255	A31	S	-1.105	A32	N	-0.804
A34	I	-0.004	A32	N	-0.624	A33	Q	0.52
A35	K	1.263	A33	Q	0.737	A34	I	-1.134
A39	N	-0.951	A34	I	-1.183	A35	K	-1.058
A40	Q	-0.218	A35	K	-1.18	A81	T	-1.471
A42	S	-1.659	A81	T	-0.946	A83	I	-0.232
A45	T	-1.344	A83	I	0.519	A85	E	-0.738
A46	K	-1.341	A85	E	0.389	A86	V	-1.02
A49	S	-1.33	A86	V	-0.923	A87	E	-1.571
A50	K	-1.427	A87	E	-1.346	A88	D	-0.887
A52	N	-1.397	A88	D	-0.72	A89	Q	-1.096
A85	E	0.563	A89	Q	-0.907	A90	K	0.187
A87	E	-1.567	A90	K	0.48	A92	E	-1.686
A88	D	-1.277	A91	E	-1.162			
A90	K	-0.688	A92	E	-0.647			
			A94	Q	-1.569			

Table 25 (Continued) The hot spots and non-hot spots with their scoring value of HSPred prediction.

p1454			p85		
position	**name**	**ΔΔG**	**position**	**name**	**ΔΔG**
A1	K	-0.978	A10	D	-1.255
A25	Q	-1.079	A11	T	-1.208
A26	F	-0.394	A43	F	-1.222
A27	H	0.22	A44	L	-0.671
A29	K	1.042	A45	T	-1.118
A30	N	-1.038	A46	K	-0.5
A31	S	-0.982	A49	S	-1.521
A32	N	-0.271	A50	K	-1.242
A33	Q	1.389	A51	L	-1.059
A34	I	-1.067	A52	N	-0.667
A35	K	-0.98	A53	D	0.114
A81	T	-1.013	A54	R	1.931
A83	I	0.582	A56	D	-0.389
A85	E	0.352	A57	S	-1.204
A86	V	-0.925	A58	R	-0.573
A87	E	-1.347	A59	R	-0.419
A88	D	-0.782	A60	S	-1.471
A89	Q	-0.905	A70	I	-1.064
A90	K	0.756	A72	K	-0.379
A92	E	-0.723	A73	N	-0.253
A94	Q	-1.591	A75	K	-1.253

APPENDIX E

International Conference

ORAL PRESENTATION

Tanchanok Wisitponchai, Piyarat Nimmanpipug, Vannajan Sanghiran Lee, and Chatchai Tayapiwatana*

"Matlab Process for Validating Amino Acids on CD4 Involving in DARPin Binding Site From ZDOCK Molecular Docking Database", The 6th Pure and Applied Chemistry Conference International Conference 2012 (PACCON 2012), Jan 11^{11}-13^{th} 2012, The Empress Chiang Mai Hotel, Chiang Mai.

*Author for correspondence

PUBLICATION

Tanchanok Wisitponchai, Piyarat Nimmanpipug, Vannajan Sanghiran Lee, and Chatchai Tayapiwatana†

"Matlab Process for Validating Amino Acids on CD4 Involving in DARPin Binding Site From ZDOCK Molecular Docking Database", Chiang Mai J. Sci. 2013; Vol.40 No.2 (APRIL 2013)

†This paper was selected to be high quality articles, so, it published in Chiang Mai J. Sci. which kindly supported the PACCON 2012 conference.

Chiang Mai J. Sci. 2012; 39(X) : 1-8
http://it.science.cmu.ac.th/ejournal/
Contributed Paper

MATLAB Process for Validating Amino Acids on CD4 Involving in DARPin Binding Site from ZDOCK Molecular Docking Database†

Tanchanok Wisitponchai [a], Piyarat Nimmanpipug [b,c], Vannajan Sanghiran Lee [c,d] and Chatchai Tayapiwatana*[a,f,g]

[a] Biomedical Engineering Center, Faculty of Engineering, Chiang Mai University, Chiang Mai 50200, Thailand.

[b] Department of Chemistry, Center of Excellence for Innovation in Chemistry and Materials Science Research Center, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand.

[c] Thailand Center of Excellence in Physics, Commission on Higher Education, Bangkok 10400, Thailand.

[d] Department of Chemistry, Faculty of Science, University of Malaya, 50603, Kuala Lumpur, Malaysia.

[f] Division of Clinical Immunology, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai 50200, Thailand.

[g] Biomedical Technology Research Unit, National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency at the Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai 50200, Thailand.

*Author for correspondence; e-mail: asimi002@hotmail.com

† *Presented at the Pure and Applied Chemistry International Conference 2012 (PACCON 2012), 11th-13th January 2012, Chiang Mai, Thailand.*

Received: 15 January 2012
Accepted: 7 April 2012

ABSTRACT

Human immunodeficiency virus (HIV) inherits the active mutation, which promotes the viral survival in human and is the main hurdle of vaccine development. The mutations on HIV surface i.e. CD4 binding site on HIV gp120, causes the escaping from the immune-surveillance but still retains the infectivity for CD4+ T cells. Formerly, Designed Ankyrin Repeat Protein (DARPin) technology has been innovated to inhibit the HIV infection. The CD4-specific DARPin specifically excludes the adhesion step of HIV to CD4 molecule and efficiently prevents HIV infection *in vitro* by competing with gp120, however, not *in vivo*. Therefore, insight into the interaction between DARPin and CD4 molecules will provide the information to improve the interaction affinity of DARPin for *in vivo* purpose. The binding activity can be modified by directly defining the key amino acid residues of CD4-specific DARPin and replacing them with other possible residues. To discover the remarkable CD4's residues, a molecular docking software i.e. ZDOCK was used to predict the complex structures between CD4 and DARPin. In this study, the candidate residues of CD4-specific DARPin were identified by the residues of CD4 molecule involving in gp120 interaction. The defined DARPin poses that interrelated with them were then extracted. The MATLAB software was implemented to generate the assigned screening criteria

for recruiting the most relevant binding residues in eleven complex-structure-candidates derived from ZDOCK. The invented data processing methodology will assist the molecular simulation researchers to rapidly and precisely define important residues.

Keywords: DARPins, CD4, HIV, Protein Docking, MATLAB.

1. INTRODUCTION

The evaluation and mutation of Human immunodeficiency virus (HIV), effect of high rates of viral replication in human host cells, are able to evade recognition by cytotoxic T lymphocytes (CTLs) [1, 2]. The entry of HIV into host cell is initiated by a binding between the viral envelope protein, gp120, and its primary receptor, CD4, on surface of T-helper lymphocytes or macrophages. The mutations on HIV surface i.e. gp120, CD4 binding site on HIV, cause the evasion of effective host-neutralizing antibodies, however, still retain the infectivity for CD4+ T cells [3]. Antiretroviral (ARV) drugs, currently available agents, are more classified by the phase of the retrovirus life cycle than direct binding with HIV particles. Several HIV drugs have been developed to hinder entry of HIV, such as maraviroc, enfuvirtide [4], and ibalizumab [5]. Drug resistance problems from virus mutations and serious side effects are the importance problem for treatments, so the discoveries for new agents are needed. One of microbicide agents for inhibiting entry of HIV has been developed, which is Designed Ankyrin Repeat Protein (DARPin) technology [6]. DARPin consists of repeat motifs and flanked constant regions. Each repeat motif, containing 33 amino acids, comprises one beta-turn and two antiparalled alpha helices [7]. Advantages of using ankyrin repeat proteins are due to their tight structure, specific binding at nanomolar range, and each repeat can contribute to target binding [8]. DARPins specific for human CD4 were reported by Schwizer *et al.* [9], *in vitro*, they are potent and highly specific to CD4 to block HIV entry. In addition, they are able to act against a wide range of virus strains and cause no effect on basis T cell function. *In vivo*, CD4-specific DARPin 57.2 efficiently binds to CD4+ cells (T helper cells, DCs, and monocytes) but has no effect on SHIV-infected rhesus macaques [10]. Therefore, insight into the interaction between DARPins and CD4 molecules may provide a clue for improving the binding activity of DARPins. To improve the binding activity, the key amino acids of CD4-specific DARPins were defined directly [11] and replaced with other possible residues. In the other hand, the replaceable amino acids on DARPin were discovered indirectly by analysis of key CD4's amino acid.

Our study aims to clarify the roles of crucial CD4's amino acid in the interaction of DARPin-CD4 complex, one idea to find replaced amino acids on DARPin, by using ZDOCK [12, 13] and MATLAB [14] processing. Protein-protein docking tool is used to predict protein complex structure. ZDOCK, an initial stage of docking, is an algorithm that optimizes desolvation, grid-based shape complementarity (GSC) and electrostatics by using fast Fourier transform (FFT) algorithm. The refinement stage, RDOCK, a small number (tens to thousands) of structures obtained in the initial stage is refined and re-ranked. The scoring from ZDOCK or RDOCK was carried out to gain the possible complex

structures. The intermolecular neighbors, binding residues, between complex structures were identified from the possible complexes. Numerous data of the binding residues were rapidly and precisely analyzed by MATLAB processing. These data are divided into five groups based on the generated criteria for the decision determining the key CD4's amino acid. Our model of key CD4's residues decision is using pair matching histogram analysis.

2. MATERIALS AND METHODS

2.1 Protein Docking

The ankyrin-CD4 docking simulations were performed by using a ZDOCK [12] and RDOCK [15] protocol in the Discovery studio (DS) 2.5. The 3D structure of CD4 specific ankyrin was generated by homology modeling [11] and the CD4 was downloaded from protein data bank (PDB code: 3CD4) [16]. The initial step, ZDOCK protocol, the docked poses were generated by 54,000 predictions around the CD4 that are possible binding regions within a small cluster radius of 6.0 Å for the RMSD cutoff, and a smaller interface cutoff of 9.0 Å. To filter the possible poses, the binding site [8] and the blocked residues of ankyrin were specified. The 2,000 docked poses with the highest scores were reported. The next step of refinement stage, the CD4 and ankyrin structures were typed by the CHARMm polarH force field. The selected docked poses with top 20 ZDOCK values were inputted into an RDOCK protocol to get complex structure.

2.2 Finding Intermolecular Neighbors

Each ankyrin-CD4 complex structure from RDOCK process, involving the gp120-CD4 complex (PDB code: 2NXY) [17], was carried out to find the binding atom neighbors. The set of neighbor atom was investigated within a distance threshold of 5.0 Å. The default parameter for the distance between the hydrogen bond donor and the acceptor was 2.5 Å and the donor proton-acceptor angles between 120-180°, were selected.

2.3 Data Transformation

Because the data from DS are in a string form, they were changed into numbering data to suit MATLAB. The data were classified into five criteria, first criterion; the number of DARPin's amino acid positions was bound to each CD4's amino acid. The second, the number of CD4's atom types in each CD4's amino acids was bound to DARPin's residues. Third criterion, the percentage of CD4's atom types in each CD4's amino acids was bound to DARPin's residues. The forth, the number of interactions in each CD4's amino acids was bound to DARPin. The last, each hydrogen-bonded CD4's amino acids were bound to DARPin's residues.

2.4 Data Selection

For each CD4's amino acid, we counted a number of times for a CD4's amino acid that was matched with the DARPin's amino acid. Then we created a histogram of the matching. In each criterion, the CD4's amino acids with the top 10 highest values in the histogram were selected to be candidates for considering key CD4's amino acids. Next the histogram values in each considered CD4's amino acids were combined and normalized by using:

$$\tilde{x} = \frac{x - \mu}{\sigma} \qquad (1)$$

, where x is histogram value of CD4's amino acid, μ is CD4's criterion mean, and σ is CD4's amino acid standard deviation.

2.5 Decision Making

We created 6 patterns of combination, i.e., patterns A, B, C, D, E and F. For each pattern, the normalized histogram values at each considered CD4's amino acid were combined. However, the criteria 1, 2, 3, 4, and 5 were used in pattern A, whereas, 1, 3, 4, and 5 were used in pattern B. For pattern C, D, E, and F, the used criteria were; 1, 2, 4, and 5; 1, 2, 3, and 4; 1, 2, and 4; and 1, 3, and 4, respectively. The combined CD4's amino acids were normalized again using equation (1). The position of the maximum normalized combined CD4's amino acid was selected for each pattern. Finally, the key CD4's residue decision, the maximum one from all patterns was selected. Because we desired only the top 3 of considered CD4's residues, the maximum was performed as 3 times.

3. RESULTS AND DISCUSSION

3.1 Characteristics of Ankyrin-CD4 Complex Structure

In 11 out of the 20 poses, the ankyrin was found to bind to the CD4 molecule on the domain 1, where the region was bound by gp120 [9]. We divided them, with the same binding region on CD4, into 3 clusters; 7 poses in the first cluster (pose 16, 21, 26, 642, 1302, and 1513); 3 poses in the second cluster (pose 1128, 1266, and 1454); and 1 pose (pose 85) in the last cluster. All clusters, the binding CD4's residues were overlapped with the binding CD4's residues that bind to gp120, as shown in figure 1.

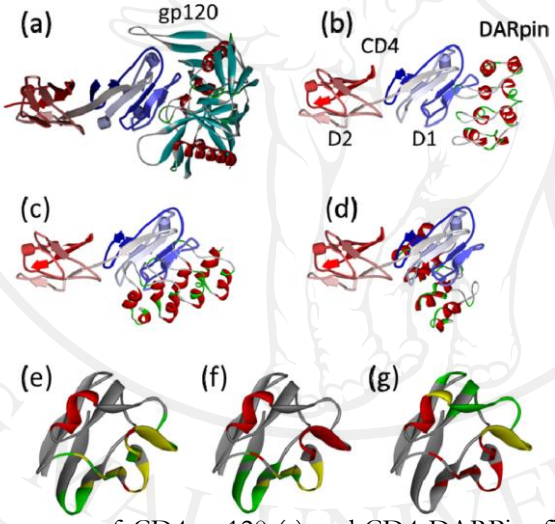

Figure 1. Complex structures of CD4-gp120 (a) and CD4-DARPin; first cluster (b), second cluster (c), and third cluster (d). Characteristics of the binding region on domain 1 of CD4; first cluster (e), second cluster (f), and third cluster (g). The red region is bound by gp120, the green region is bound by DARPin, and the yellow region is bound by both DARPin and gp120.

3.2 Data Analysis

In every pose, the values in each criterion, excluding the fifth criterion, were counted to be considered as CD4's amino acids, as shown in figure 2. The considered CD4's amino acids to be key residues in each poses were shown in table 1. In cluster 1, 7 poses, there were 11 considered CD4's amino acids that were the same 100%; these amino acids were Q25, H27, Q33, I34, K35, N39, Q40, G41, P48, E85, E87, and K90. In cluster 2, 3 poses, there were 10 considered CD4's amino acids that were the same 100%. These amino acids were H27, K29, S31, N32, Q33, E85, E87, D88, K90,

Table 1. The considered CD4's amino acids that were a combination of top ten in each criterion.

Pose	position of considered CD4's amino acids
16	25, 27, 33, 34, 35, 39, 40, 41, 48, 85, 87, 90
21	25, 27, 29, 33, 34, 35, 39, 40, 41, 48, 85, 87, 90
26	25, 27, 29, 33, 34, 35, 39, 40, 41, 48, 85, 90
241	25, 27, 33, 34, 35, 39, 40, 41, 47, 48, 85, 87, 90
642	25, 27, 29, 33, 34, 35, 39, 40, 41, 48, 85, 87, 88, 90
1302	25, 27, 29, 32, 33, 34, 35, 39, 40, 41, 48, 85, 90
1513	25, 27, 29, 32, 33, 34, 35, 39, 40, 41, 48, 85, 87, 90
1128	1, 25, 27, 29, 31, 32, 33, 85, 87, 88, 90, 92
1266	27, 29, 31, 32, 33, 35, 81, 85, 87, 88, 90, 92
1454	25, 27, 29, 31, 32, 33, 35, 85, 87, 88, 90, 92
85	9, 44, 45, 46, 52, 53, 54, 56, 59, 60, 72, 73

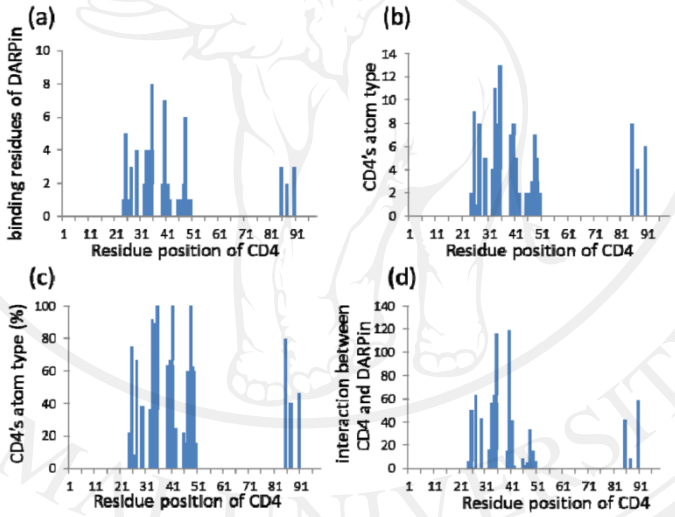

Figure 2. Example of pose 26, one of cluster 1, the binding data are shown in four criteria; the first (a), the second (b), the third (c), and fourth (d) criterion.

and E92. The identical CD4's residues considered between cluster 1 and 2 were H27, Q33, E85, E87, and K90. These results may coincide with the key amino acids of CD4.

In each pose, the normalized histogram values were carried out to detect the top 3 of the key CD4's residues by using maximum detection. The top 3 residues were shown in table 2. In cluster 1, K35 was found the most key CD4's residues, with 100% probability, followed by Q40, Q25, Q33, and P48 with probability of 85.7%, 57.1%, 42.9%, and 14.3 respectively. The results implied that K35 is the most probable CD4's amino acid to be considered as the key residue of CD4-DARPin interaction because of the above reason. Similarly, it had 63.6% probability when we considered throughout 11 poses. Moreover, the possibility of the first residue was 71.4%. Q40 and Q25 was the secondary and tertiary important residue in the cluster 1. In cluster 2, N32, Q33, and K90 had 100 % probability as the top 3 key residues. When we considered possibility of the first key residues, Q33 and K90 had

probability of 66.7% and 33.3%, respectively. Furthermore, Q33 was found in cluster 1 with a probability of 42.9%, and at 54.5% when considered throughout the 11 poses. Hence, in cluster 2, Q33 was the most probable key residue, followed by K90 and N32, respectively. In cluster 3, R59, S46, and N52 were the first, second, and third key residues for binding with DARPin. Therefore, the key amino acids in cluster 1 were K35 (as shown in figure 3) Q40, and Q25; cluster 2 were Q33, N32, and K90; cluster 3 were R59, K46, and N52. All key CD4's amino acids were located on the binding site of gp120 specific CD4 within a 5 Åradius,except for K90 and D93 (data not shown). According to the key amino acids, K35 and R59, were found to be parts of the critical residues on CD4 that gp120 recognition, which were Lys29, Lys35, Phe43, Leu44, Lys46, Gly47 and Arg59; these were studied by biochemical mutagenesis (Phe43 and Gly47) [18] and compiled from mutagenesis studies by Ryu et.al [19]. Specific mutation experiments to confirm the theoretical investigation of key CD4's amino acids will be undertaken and observed for the interactions between mutated CD4 and DARPin.

Table 2. The top 3 key CD4's amino acids bound to DARPin

Pose	Key CD4 amino acid		
	1st	2nd	3rd
16	35	25	40
21	40	35	25
26	35	40	33
241	40	35	25
642	35	40	25
1302	35	48	33
1513	35	33	40
1128	33	90	32
1266	33	90	32
1454	90	32	33
85	59	46	52

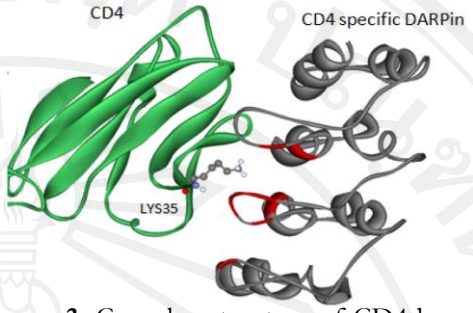

Figure 3. Complex structure of CD4 bound to DARPin. The CD4:LYS35 is one of validating amino acid on CD4 related with DARPin binding site. The relative binding residues of CD4-specific DARPin are marked in red color.

4. CONCLUSIONS

The binding site pattern from our simulation of CD4-ankyrin docking was related with Schweizer *et al.* [9] report. Although we obtained different binding patterns, the CD4-DARPin binding regions of all three patterns overlap with CD4-gp120 interacting region. Moreover, some key amino acids from our model are part of the crucial CD4 that bind to gp120 [18, 19]. They are the candidates for finding the relevant amino acids of ankyrin, which will be mutated further to enhance the binding affinities. Also, from this study, we have shown that effective data analysis for a very large dataset can be done successfully in MATLAB.

ACKNOWLEDGEMENTS

This work was supported by the Thailand Research Fund (TRF), the National Centre

for Genetic Engineering and Biotechnology (BIOTEC) of the National Science and Technology Development Agency (NSTDA), the National Electronics and Computer Technology (NECTEC) for Discovery software. We would like to thank Associate Professor Nipon Theera-Umpon and Assistant Professor Sansanee Auephanwiriyakul for consultations on the data analysis.

REFERENCES

[1] Lieberman J., Defying Death-HIV Mutation to Evade Cytotoxic T Lymphocytes, *N. Engl. J. Med.*, 2002; **347:** 1203-1204.

[2] Troyer R.M., McNevin J., Liu Y., Zhang S.C., Krizan R.W., Abraha A., and et al., Variable Fitness Impact of HIV-1 Escape Mutations to Cytotoxic T Lymphocyte (CTL) Response, *PLoS Pathog.,* 2009; **5**: e1000365.

[3] Bolmstedt A., Hemming A., Flodby P., Berntsson P., Travis B., Lin J.P., and et al., Effects of Mutations in Glycosylation Sites and Disulphide Bonds on Processing, CD4-Binding and Fusion Activity of Human Immunodeficiency Virus Envelope Glycoproteins, *J. Gen. Virol.,* 1991; **72 (Pt 6):** 1269-1277.

[4] Pugach P., Ketas T.J., Michael E., and Moore J.P., Neutralizing Antibody and Anti-Retroviral Drug Sensitivities of HIV-1 Isolates Resistant to Small Molecule CCR5 Inhibitors, *Virology*, 2008; **377:** 401-407.

[5] Bruno C.J., and Jacobson J.M., Ibalizumab: an Anti-CD4 Monoclonal Antibody for the Treatment of HIV-1 Infection, *J. Antimicrob. Chemother.,* 2010; **65:** 1839-1841.

[6] Binz H.K., Stumpp M.T., Forrer P., Amstutz P., and Pluckthun A., Designing Repeat Proteins: Well-Expressed, Soluble and Stable Proteins from Combinatorial Libraries of Consensus Ankyrin Repeat Proteins, *J. Mol. Biol.*, 2003; **332:** 489-503.

[7] Sedgwick S.G., and Smerdon S.J., The Ankyrin Repeat: a Diversity of Interactions on a Common Structural Framework, *Trends Biochem. Sci.,* 1999; **24:** 311-316.

[8] Zahnd C., Wyler E., Schwenk J.M., Steiner D., Lawrence M.C., McKern N.M., and et al., A Designed Ankyrin Repeat Protein Evolved to Picomolar Affinity to Her2, *J. Mol. Biol.,* 2007; **369:** 1015-1028.

[9] Schweizer A., Rusert P., Berlinger L., Ruprecht C.R., Mann A., Corthesy S., and et al., CD4-Specific Designed Ankyrin Repeat Proteins are Novel Potent HIV Entry Inhibitors with Unique Characteristics, *PLoS Pathog.,* 2008; **4:** e1000109.

[10] Pugach P., Krarup A., Gettie A., Kuroda M., Blanchard J., Piatak M. and et al., *In Vivo* Binding and Retention of CD4-Specific DARPin 57.2 in Macaques, *PLoS One*, 2010; **5:** e12455.

[11] Nimmanpipug P., Khampa C., Lee V.S., Nangola S., and Tayapiwatana C., Identification of Amino Acid Residues of a Designed Ankyrin Repeat Protein Potentially Involved in Intermolecular Interactions with CD4: Analysis by Molecular Dynamics Simulations, *J. Mol. Graph. Model.,* 2011; **31:** 65-75.

[12] Chen R. and Weng Z., A Novel Shape Complementarity Scoring Function for Protein-Protein Docking, *Proteins,* 2003; **51:** 397-408.

[13] Chen R., Li L., and Weng Z., ZDOCK: an Initial-Stage Protein-Docking Algorithm, *Proteins,* 2003; **52:** 80-87.

[14] Sobie E.A., An Introduction to MATLAB, *Sci. Signal.,* 2011; **4:** tr7.

[15] Li L., Chen R., and Weng Z., RDOCK: Refinement of Rigid-Body Protein Docking Predictions, *Proteins,* 2003; **53:** 693-707.

[16] Garrett T.P., Wang J., Yan Y., Liu J., and Harrison S.C., Refinement and Analysis of the Structure of the First Two Domains of Human CD4, *J. Mol. Biol.,* 1993; **234:** 763-778.

[17] Zhou T., Xu L., Dey B., Hessell A.J., Van R.D., Xiang S.H., and et al., Structural Definition of a Conserved Neutralization Epitope on HIV-1 gp120*, Nature,* 2007; **445:** 732-737.

[18] Wu H, Myszka DG, Tendian SW, Brouillette CG, Sweet RW, Chaiken IM, and et al., Kinetic and Structural Analysis of Mutant CD4 Receptors that are Defective in HIV gp120 Binding, *Proc. Natl. Acad. Sci. U. S. A.,* 1996; **93:** 15030-15035.

[19] Ryu S.E., Truneh A., Sweet R.W., and Hendrickson W.A., Structures of an HIV and MHC Binding Fragment from Human CD4 as Refined in Two Crystal Lattices, *Structure,* 1994; **2:** 59-74.

VITA

Name Miss Tanchanok Wisitponchai

Date of Birth 09 November 1987

Educational Background High School Certificate Holder from Satree Rachinuthit School, Udonthani (2005)

Bachelor of Science (Medical Technology, First class Honor) degree holder from Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand (2009)